PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G06F 1/00 — A1

(11) International Publication Number: WO 98/53387

(43) International Publication Date: 26 November 1998 (26.11.98)

(21) International Application Number: PCT/US98/09969

(22) International Filing Date: 18 May 1998 (18.05.98)

(30) Priority Data:
60/047,235 20 May 1997 (20.05.97) US

(71) Applicant: AMERICA ONLINE, INC. [US/US]; 22000 AOL Way, Dulles, VA 20166 (US).

(72) Inventors: MORRIS, Harry, W.; Apartment F, 1719 Ascot Way, Reston, VA 20190 (US). BOSCO, Eric; 2430 13th Court N., Arlington, VA 22201 (US). LIPPKE, David, Lowell; 1441 Kingsvale Circle, Herndon, VA 20170 (US). STEELE, Colin, Anthony; 635 Cobbler Terrace, Leesburg, VA 20175 (US).

(74) Agent: PHILLIPS, John, C.; Fish & Richardson P.C., 601 Thirteenth Street, N.W., Washington, DC 20005 (US).

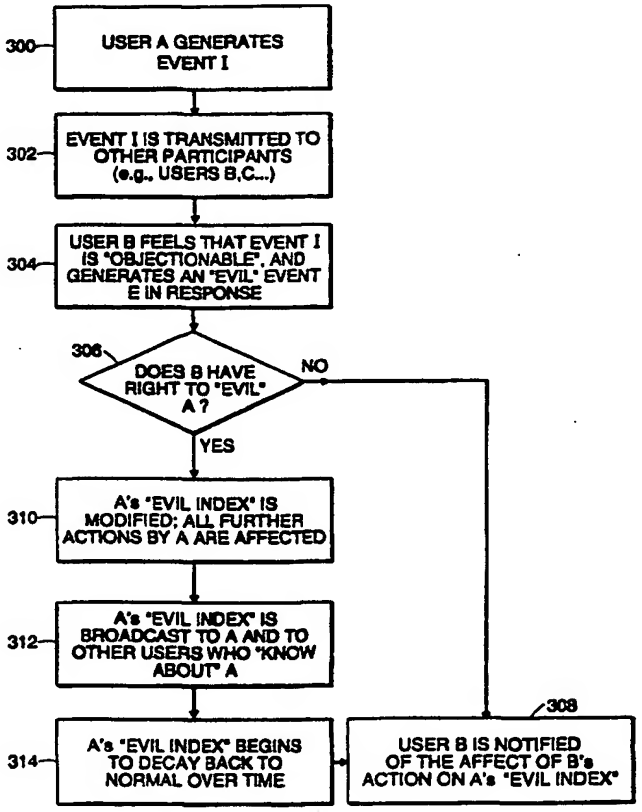
(81) Designated States: AU, BR, CA, JP, MX, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: SELF–POLICING, RATE LIMITING ONLINE FORUMS

(57) Abstract

A method and means for self–policing and automatically rate–limiting multiple–user online forums. The preferred embodiment of the invention includes a set of rules that permit users to censure other users. A censured user has one or more "privileges" (which may include access to the online computer system or the effective message rate of a communications connection) taken away or diminished; the privilege is gradually restored if the censure user behaves. The censuring participant and other pertinent participants are notified of the effect of the censuring on the censured user. In another aspect of the invention, the online computer system automatically tracks the rate at which a user sends certain types of messages, and can message rate limit a user who uses too many system resources by sending a large number of messages in rapid succession.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

SELF-POLICING, RATE LIMITING ONLINE FORUMS

BACKGROUND

This application is a continuation of provisional application number 60/047,235 filed May 20, 1997.

Technical Field

This invention relates to distributed computer services, particularly computer services having online forums.

Background Information

An online forum is a communications interchange in which people may communicate with others through successive electronic transmissions between respective computer systems. An online forum, or any other type of distributed computer services, may be implemented on a distributed computer system such as that shown in FIG. 1. Forum participants (equivalently, users of the computer services) typically are scattered across a large geographical area and communicate with one or more central server systems 100 through respective client systems 102 (*e.g.*, a personal or laptop computer). In practice, the server system 100 typically will not be a single monolithic entity but rather will be a network of interconnected server computers, possibly physically dispersed from each other, each dedicated to its own set of duties and/or to a particular geographical region. In such a case, the individual servers are interconnected by a network of communication links, in known fashion. One such server system is "America Online" from America Online Incorporated of Virginia.

Each client system 102 runs client software that allows it to communicate in a meaningful manner with corresponding software running on the server system 100.

The client systems 102 communicate with the server system 100 through various channels, such as a modem 104 connected to a telephone line 106 or a direct Internet connection using a transfer protocol such as TCP/IP (Transfer Control Protocol/Internet Protocol). The server system 100 is responsible for receiving input from the client systems 102, manipulating the collective body of input information (and possibly information from other sources) into a useful format, and retransmitting the formatted information back to one or more clients 102 for output on an output device, such as a display screen.

Referring to FIG. 2, one type of forum is a "chat room" 200, in which the various participants 204 (*e.g.*, "Allens9," "JOSHUAALEX," *etc.*) may enter text which appears in a scrolling text window 202 on each participant's computer display screen. In the example in FIG. 2, the chat room 200 has 22 participants whose identities (or "screen names") are listed in a scrolling window 210. A participant 204 may respond to the comment of another participant 204 by entering a line of text in an edit box 206 and activating (*e.g.*, by clicking with a pointer device, such as a mouse) a SEND button 208. In response, the text in the scrolling text window 202 scrolls upwards and the newly entered line of text is displayed at the bottom of the scrolling text window 202. In the illustrated example, the last participant to enter a comment was JOSHUAALEX, who typed "TEXAS."

The chat room 200 shown in FIG. 2 is "public", meaning that it has multiple participants who were placed in the chat room by the computer-service provider and who most likely never have met or conversed with one another before. A comment by a participant in a public forum may be seen by all of the participants of the chat room. If a participant desires some privacy, that participant may

"open" and enter a "private" chat room (for example, by clicking on a SETUP button 212), and thereafter invite one or more other participants to enter the private chat room. Once in a private forum, participants may communicate with one another without fear that uninvited participants will be able to see their comments.

When a participant in a forum, whether public or private, makes a comment that others in the forum regard as offensive, in poor taste, wildly incorrect, or otherwise objectionable, the offending participant most likely will be "flamed" by one or more of the other participants. A "flame" is a reprimand or other stringent response directed at the offending party. One purpose behind flaming another participant is to dissuade the offender, through embarrassment or intimidation, from making further objectionable comments. In this manner, if the offending user chooses to curb his or her behavior in response to the flaming, a forum may be crudely regulated or "policed" by the forum's participants. However, the offending participant may continue to behave in an objectionable manner. Further, a participant who overly "flames" other participants may also be objectionable. Accordingly, participant policing of forums does not always work well. In such cases, offended participants may drop out of "flame-filled" forums, and/or the online service must devote resources to actively police problematic participants.

Other objectionable behavior includes sending one or more messages to "spoof" other users as to the sender's identity in order to try to get confidential information (*e.g.*, credit card numbers or passwords) sent in response (sometimes called "password fishing").

Another problem that can arise in online systems is "resource hogging", where a participant uses features such as broadcast or multi-cast messaging to send a large number of messages to other users in a short period of time (sometimes called "spamming"). Such resource hogging deprives other users of server resources, and can slow an online system response time to undesirable levels.

Accordingly, the inventor has determined that there is a need for a better way to police recalcitrant participants in online forums and to reduce spamming. The present invention provides a method and means for accomplishing this goal.

SUMMARY

The invention provides a method and means for self-policing and automatically rate-limiting multiple-user online forums. The preferred embodiment of the invention includes a set of rules that permit users to censure other users. A censured user has one or more "privileges" (which may include access to the online computer system or the effective message rate of a communications connection) taken away or diminished; the privilege is gradually restored if the censured user behaves. The censuring participant and other pertinent participants are notified of the effect of the censuring on the censured user. In another aspect of the invention, the online computer system automatically tracks the rate at which a user sends certain types of messages, and can message rate limit a user who uses too many system resources by sending a large number of messages in rapid succession. The amount of rate limiting may be a function of the amount of censure that has been applied to the user being limited.

Advantages of this invention may include one or more of the following. The techniques described here enable a multiple-user online service (*e.g.*, a chat room or other forum) to be self-policing. Access to the computer service for a particular user is automatically regulated by the computer based on input from other users concerning the conduct of the user under consideration, or based on the message rate of a user. Users of a computer-based system have the ability to sanction a misbehaving user and thereby have the offending user's access to the system denied or curtailed. Unlike the conventional "flaming" approach to policing, which typically fails because it relies on the misbehaving user curbing his or her own behavior or results in retaliatory "flaming", the policing techniques described here are based on predetermined rules and operate automatically in response to votes cast by other users of the computer system. The offending user has no choice in the matter once he or she chooses to misbehave. Because the operation of these policing techniques are automatic, the computer-based service provider need not expend personnel time and resources to police public forums, direct user communications, and the like.

The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, objects, and advantages of the invention will be apparent from the description and drawings, and from the claims.

DESCRIPTION OF DRAWINGS

FIG. 1 shows a prior art distributed computer system of the type used for providing online computer services.

FIG. 2 is a screen shot showing an example of a prior art online computer forum.

FIG. 3 is a flowchart of a basic embodiment of the self-policing aspect of the invention.

FIG. 4 is a flowchart of a basic embodiment of the rate-limiting aspect of the invention.

Like reference numbers and designations in the various drawings indicate like elements.

DETAILED DESCRIPTION

The prospect of millions of concurrent users of an online computer system puts great pressure on the ability of the system provider to police abusive participants. Accordingly, the invention provides a self-policing online environment that makes participants responsible for their behavior. That is, other participants can express an opinion about that behavior, and have that opinion affect the offending user in a negative way and be "visible" to other participants. Providing such a self-policing user feedback mechanism lowers the cost of providing online services to users and provides a more "user-friendly" environment for participants.

In another aspect of the invention, the online computer system automatically tracks the rate at which a user sends certain types of messages, and can "rate limit" a user who "hogs" too many system resources by sending a large number of messages (or messages of selected types) in rapid succession.

Participant Self-Policing

FIG. 3 is a flowchart of a basic embodiment of the self-policing aspect of the invention. Initially, user A generates an event I, such as a message (STEP 300). Event I is transmitted to some number of other participants, such as users B and C in a public forum on an online

computer system (STEP 302). However, event I may be a direct communication between user A to user B, such as by use of the Instant Message™ feature of the America Online computer service. User B may feel that event I is "objectionable" or "evil" (a subjective decision by B), and therefore generates a "vote" against the behavior of user A by sending a special type of response message - an "evil" event E (STEP 304). In the preferred embodiment, a user cannot "evil" another participant except in response to a message from the participant, or otherwise in cases where the participant's actions effect the user doing the "eviling." In order to reduce retaliation, users cannot "evil" other participants directly in response to being "eviled".

The online system receives event E and checks a database to see whether user B has the right to "evil" user A (STEP 306). For example, this step can limit users who themselves have been restricted from "eviling" other users.

If user B does not have the right to "evil" user A, then user B is notified of the affect of user B's action on user A's "evil index" - a value of how restricted user A is with respect to use of resources on the online system (STEP 308). In this case, user B's action has no affect on user A.

If user B does have the right to "evil" user A, then user A's "evil index" is modified by a suitable amount, which will affect further actions by user A (STEP 310), as described below. The amount of modification can be based on the past behavior of user A, or of users A and B. The amount of modification may also depend on the type of "eviling" asserted by user B. For example, user B may wish to assert an "eviling" event E anonymously rather

than be identified. Anonymous "eviling" may be accorded lesser weight. That is, an anonymous "eviling" response may count as fewer evil "votes" than it would if the eviling user's identity was revealed. In one embodiment, an "eviling" user may set up an automatic "eviling" response to all communications from another specific user. In this case, the effect of the "eviling", which is intended to discourage communication from the specific user, is given very low weight (equivalently, counts as very few evil "votes") because of its automatic invocation.

User A's "evil index" is "broadcast" to other users who "know about" user A (STEP 312). For example, such users might be all of the participants in a private chat room, or all visitors to a public chat room. In one embodiment, a modified "evil index" is reported to each user in a current chat room's user list and to each user who has the "eviled" users on his or her "buddy list" (described in co-pending U.S. Patent Application Serial No. 08/803,692, filed 2/24/1997, entitled "User Definable On-line Co-user Lists". Thus, an "eviled" user is publicly chastised.

Thereafter, user A's "evil index" begins to gradually "decay" back to normal over time (STEP 314). Such decay may be linear, exponential, step-wise, or some other function. Meanwhile, user B is notified of the affect of user B's action on user A's "evil index" (STEP 308). In this case, user B's action has had an affect on user A.

Some of the steps noted above may be done in different order without substantially changing the effect of the process. For example, STEPS 312, and 314, and 308 may be performed in different order.

A basic "penalty" for having a non-normal "evil index" is denial of access to a forum or the online service until the user's "evil index" has decayed back to normal. In a more refined embodiment, a user's "evil index" affects a rate limit which governs a user's ability to send (and/or receive) messages. This feature allows other participants to "evil" a user who "flames" or "spams" them, and thus reduce the rate at which the recalcitrant user can send and/or receive messages. A description of rate limiting is set forth below.

A server database (which may be centralized or distributed) stores a user's "evil index". A user's "evil index" can be maintained in a user-record as a global total, or by forum, or both. The value of each "evil index" can be used to control the user's ability to log on to the online computer system or access selected forums, and/or the effective rate of message or file transmissions.

Other rules may applied in different embodiments. For example, the following rules can be implemented:

A user must be in a forum (*e.g.*, a chat room, but including direct communication between users, such as the Instant Message™ feature of the America Online computer service) for a specified amount of time before being allowed to "evil" another user in that forum; this reduces "hit-and-run flaming". The minimum amount of time may vary from forum to forum, and from user to user (*e.g.*, a forum "sysop" may be immune to "eviling"). A user's user-record in the server database thus would record a time-of-entry for a forum. For example, a user's time-of-entry to a forum would be compared to the current time in STEP 306 of FIG. 3 to

determine if the user had the right to "evil" another participant.

A user must be in a forum for a specified amount of time before being allowed to "evil" another user who has been in that forum for a longer period of time. For example, a user's time-of-entry to a forum would be compared to the time-of-entry of another participant in STEP 306 of FIG. 3 to determine if the user had the right to "evil" that other participant. The specified amount of time may vary from forum to forum, and from user to user.

A user's eviling response may be accorded increased weight (equivalently, counted as extra "evil" votes) based on the "seniority" of the "eviling" user. Each additional unit of time spent in a forum could enhance a user's "seniority," thereby allowing long term user's more "eviling" power than new-comers. That is, being "eviled" by an "old-timer" can have a greater than normal affect on modification of a chastised user's "evil index" in STEP 310 of FIG. 3. A user's user-record in the server database thus would record a total-time-of-access for each forum, which could be a "lifetime" total (such that users who had belonged to the online computer service for longer periods would have greater "eviling" power than relatively new subscribers to the service) , a running total for a selected period of time, or a session total. The time period for accumulating extra votes and the "eviling" value of extra votes may vary from forum to forum, and from user to user.

A user may be limited in the number of "eviling" votes that can be cast in any one online session or time period (*e.g.*, a day or week). A user's user-record in the server database thus would record the number of "eviling" votes cast globally or by forum. For example, a user's number of previous "eviling" votes cast could be examined in STEP 306 of FIG. 3 to determine if the user had the right to "evil" another participant.

The amount of modification of a user's "evil index" in STEP 310 of FIG. 3 after being "eviled" is preferably non-linear, so that the first few times a user is "eviled" has little effect (and possibly no effect until some threshold level of accumulated "eviling" votes from several participants are asserted against the user) on the user's "evil index" (everyone makes a few mistakes). However, for subsequent times that a user is "eviled", the user's "evil index" preferably is modified to a much greater degree, increasing the penalty for recidivism. The determination of how much to modify a user's "evil index" can be based on the user's global "evil index" (*i.e.*, a "rap sheet" evaluation) or forum "evil index" (*i.e.*, a "fresh start" approach). A user's user-record in the server database would thus record the number of times the user has been "eviled", either globally or by forum, for example. The recorded number might be a "lifetime" total or a running total for a selected period of time. The amount of "evil index" modification may also vary from forum to forum, and from user to user.

As an example, each user may be given an initial "evil index" of 100. A first "offense" may reduce that value to 95; a second "offense" reduces the value to 85; third and subsequent offenses reduce the current value by 15 units. Alternatively, each user is given an initial "evil index" of 0 and "eviling" increases that value. If a range of 100 is used, a user's "evil index" may be regarded as a "percentage of evil", up to 100% "evil".

The decay rate for returning a user's modified "evil index" to normal can vary from forum to forum. For example, the decay in chat rooms (where most flaming occurs) may be less than in other forums. Using the example from immediately above, the user's "evil index" may "decay" back to "normal" at the rate of 2 units per hour in a chat room, but at the rate of 5 units per hour in other forums.

A user may query a forum to determine the rate of "eviling" within the forum as a measure of how much members of the forum "evil" one another. A high rate may indicate that one or more members are misbehaving. The system keeps track of all users participating in a forum, so the current "eviling" rate is a time-weighted average of the number of "eviling" votes cast by the current participants. In an alternative embodiment, a user may query a forum to determine the accumulated value of the "evil index" of all current users. This measures how often the users have misbehaved in the past, and can serve as an estimate of the likelihood that the current users will misbehave in the future. The system keeps track of all users participating in a forum, so the total "evil

index" is the sum of the pertinent forum "evil index" for each participant.

In some embodiments, a user who has been "eviled" has a lessened ability to "evil" other users (thus reducing retaliation). However, some online systems implement message types, such as broadcast or multi-cast messages or self-repeating messages, that are more frequently used than other message types to flame or spam other participants. In some embodiments, it may be desirable to allow all users (even those with a modified "evil index", and thus possibly restricted in usage rights) to "evil" the originator of such message types. Such a determination would be made in STEP 306 of FIG. 3 by examining the message type of event I sent by the originating user. The amount of modification of such an originator's "evil index" can be set to be greater than for flaming to discourage use of such message types for spamming.

When a user's "evil index" is modified, the user is notified that privileges, such as rate of messaging, have been limited. In one embodiment, a message is sent from the server that contains the user's current "evil index" to the "eviled" user, and optionally includes the current rate of decay back to normal. Such information allows a wide range of feedback to be presented to the user about his or her ability to interact with the online computer system. For example, a graphical "power meter" or "power bar graph" can be used to indicate the "evil index" of the user. For instance, a color-coded bar graph can be divided into (1) a green zone to represent that the user's "evil index" is normal; (2) a yellow zone to represent that the user's "evil index" has been modified slightly; (3) a red zone to represent that the user's

"evil index" has been modified significantly; and (4) a black zone to represent that access or message privileges have been suspended for a time. However, other methods of informing a recalcitrant user of his or her "evil index" can be used.

In one embodiment, a user can "rehabilitate" his or her "evil index" by visiting advertisements displayed on the online computer system; "visits" can be determined by requiring the user to "click" on an advertisement. The user is rewarded with more "power" by adjusting the value of the user's "evil index" more towards normal.

Automatic Rate Limiting

In one embodiment, both input and output messaging rates of a user are limited based on the behavior of the user and/or available system resources. Such rate limiting can stop malicious users and errant client computers from consuming more than a fair share of online system resources. However, preferably the rate limiting system is weighted to tolerate brief bursts of messaging activity while penalizing unacceptably large rates of messaging. Rate limiting can also limit aggregate input to a server to a level at which the system is reasonably loaded under normal conditions. Rate limiting can also be combined with "eviling" by automatically adjusting a users rate limit parameters based on their "evil index."

In one embodiment, input rate limiting - from user to server - is conducted on a per user connection basis, and within a user connection on a per-message type bases (that is, rate limits for different types of messages may be set to different values). In one embodiment, rate limiting for a user is achieved in accordance with the following algorithm:

(1) Define *A* as the running average of inter-message time gaps for the last *N* messages of selected types that the user has attempted to send; a system selected value *I* is used as the first value for *A*. Calculation of *A* can be done, for example, on a forum basis (accounting only for messages sent in the current forum), session basis (accounting for messages sent in the user's current online session), or message-count basis (accounting for the last *N* messages sent at any time in any forum).

(2) If *A* is below a warning threshold *W* (indicating that the user is approaching a point of sending messages too frequently), when the user attempts to send a message, send the user a warning message but transmit the user's message.

(3) If *A* is below a rate limit threshold *R* (indicating that the user is sending messages too frequently) when the user attempts to send a message, send the user a warning message and drop the user's message.

(4) Repeat the above steps until *A* rises above a clear threshold *C* (indicating that the user is not sending messages too frequently), at which time the rate limiting condition is considered cleared.

(5) If at any time *A* drops below a disconnect threshold *D*, disconnect the user.

FIG. 4 is a flowchart of a basic embodiment of the rate-limiting aspect of the invention, showing a slightly different order for the steps described above. A user attempts to send a message (STEP 400). Average *A* is computed (STEP 402). If *A* is less than a disconnect threshold *D* (STEP 404), the user is disconnected (STEP 406). Otherwise, if *A* is less than a rate limit threshold *R* (STEP 408), the message is dropped, the user is warned

that the rate limited has been exceeded, and the user is flagged as "rate limited" (STEP 410).

Otherwise, if *A* is less than a warning threshold *W* (STEP 412), a determination is made as to whether the user is rate limited (STEP 414). If not, the message is sent, but the user is warned that the rate limit is being approached (STEP 416). Otherwise, the message is dropped and the user is warned that the rate limited has been exceeded (STEP 418).

If *A* is not less than the warning threshold *W* (STEP 412), and *A* is less than a clear threshold *C* (STEP 420), a determination is made as to whether the user is rate limited (STEP 421). If not, the message is sent (STEP 423). Otherwise, the message is dropped and the user is warned that the rate limited has been exceeded (STEP 418).

Finally, if *A* is not less than the clear threshold *C* (STEP 420), the rate limit flag for the user is cleared (STEP 422) and the message is sent (STEP 423).

The rate limiting algorithm supports several "tunable" parameters:

- The running average of inter-message time gaps - *A*
- The number of message receptions over which *A* is calculated - *N*
- An initial average - *I*
- A clear threshold - *C*
- A warning threshold - *W*
- A rate limit threshold - *R*
- A disconnect threshold - *D*

In one embodiment, the values for *C*, *W*, *R*, and *D* are selected such that $C > W > R > D$. The initial average

rate *I* can be weighted to increase the algorithm's tolerance of bursts of activity, such as "chatty startups" when a user joins a forum and sends several messages in succession. The threshold rates can be set globally for all user's, or "tuned" for each user.

In one embodiment, the difference between the clear threshold *C* and the rate limit threshold *R* can be "tuned" by the online computer system to alter the interval between commencement of rate limiting for a user and the resumption of normal activity. The difference between *C* and *R* may be automatically increased, for example, if the user sends excessive "spamming" type messages at too great a rate. Further, the threshold values *C*, *W*, *R*, and *D* can be dynamically modified by the online system as a way of limiting system resource usage if too many users are simultaneously using the system.

As noted above, when a user's message rate is limited, the user is notified. In one embodiment, a message is sent from the server to a rate limited user that contains values for each of the parameters outlined above, and another message is sent the next time that the server will accept messages from the user without restriction. Such information allows a wide range of feedback to be presented to the user about his or her ability to interact with the online computer system. For example, a graphical "power meter" or "power bar graph" can be used to indicate the "health" or "power" of the user. For instance, a color-coded bar graph can be divided into (1) a green zone to represent that no rate limiting is in effect; (2) a yellow zone to represent that the user's message rate is near the point where rate limiting would take effect; (3) a red zone to represent that message rate limiting is in effect; and (4) a black zone to represent that access privileges have been suspended for

a time. However, other methods can be used to inform a recalcitrant user of his or her message rate limit status.

In one embodiment, a user can "rehabilitate" his or her message rate limit status by visiting advertisements displayed on the online computer system. The user is rewarded with more "power" by adjusting the parameters.

In one embodiment, output rate limiting - from server to user - is performed in a similar manner, with the following differences:

6) The system server audits the rate of outgoing error messages of selected types (*e.g.*, RATE_TO_HOST), and uses the above rate limiting algorithm with $R=D$, to prevent excessive load due to handling users that are generating rate errors.

7) The system server can limit the rate of outgoing messages of selected types (*e.g.*, those types that may be used for spamming) using the above rate limiting algorithm when $A < R$. In this case, a notice message can be sent to the user, to the user's intended recipient, and to a server storing the user's "evil index". The intended recipient can "evil" the sending user, or the system server can automatically modify the user's "evil index" based on the sending of excessive "spamming" type messages at too great a rate.

The algorithms described above prevent abuse of online system resources without causing undue "pain" to users, including "power users" and users who are subject to delays and input/output timing changes by non-ideal networks (which can cause otherwise normal user activity to appear to be unacceptable activity).

Rate Limiting and "Eviling"

The concepts of automatic rate limiting and modification of a user's "evil index" can be combined to provide a refined self-policing, automatic rate limiting system that can regulate an "eviled" user's ability to participate in forums without requiring total exclusion from the forum. Further, a user's current global or forum "evil index" can be used to modify the parameters used to effect automatic (*i.e.*, "non-eviled" basis) rate limiting for the user. For example, a user's ability to interact in a forum can be rate limited by modifying the values for *A*, *C*, *W*, *R*, and/or *D* as a function of whether the user's "evil index" has been modified, or as a function of the current value of the user's "evil index". Conversely, the amount of modification of a user's "evil index" can be a function of the current value of *A*. Thus, conceptually, a user's permitted message rate R_p is a function of the user's "evil index" *EI* plus the user's rate of attempted message activity *A*: $R_p = f(EI) + g(A)$

Conceptually, each user's user-record thus may look like the following table:

Forum ID	"evil index"	decay rate	# times "eviled"	# times "eviled" others	time-of-entry	total-time of-access	*A*	*C*	*W*	*R*	*D*
Global value											
Forum 1 value											
Forum 2 value											
etc.											

Separate "lifetime" and "session" records may be kept where useful. Of course, other or different information may be recorded for each user, and other ways of organizing such data may be used. Further, users may be

assigned a "class" designation (such as "sysop" (system operator), corporate users, paying members, non-paying members, etc.) which can be given weight in making a determination of right to "evil" or absolute or relative immunity from being "eviled" in STEP 306 of FIG. 3.

In setting rate parameter values for a user, a global rate table can be used which is indexed by a user's "evil index"; that is, message rate parameters are a function of the user's global or forum "evil index". A conceptual example of one such table might be the following (ΔA represents an optional adjustment to be applied to the calculated value of A):

"evil index" value	ΔA	C	W	R	D
0-20					
21-40					
41-60					
61-80					
81-90					
91-95					
96-100					

The values for A, C, W, R, and D in this table can be defined globally or by forum. If by forum, the values can be defined in absolute terms or as relative offsets to a global table. Multiple tables of this form also can be defined, indexed by message type, so that misuse of certain types of messages are punished more than misuse of other message types.

Alternatively, a non-tabular implementation could be used instead of the global rate table described above. For example, a user's rate parameter values (ΔA, C, W, R, D)

could be defined as a function describing a relation between these quantities.

Implementation

The methods and mechanisms described here are not limited to any particular hardware or software configuration, but rather they may find applicability in any computing or processing environment used in connection with online computer services.

The invention may be implemented in hardware or software, or a combination of both. However, preferably, the invention is implemented in computer programs executing on programmable computers each comprising at least one processor, at least one data storage system (including volatile and non-volatile memory and/or storage elements), at least one input device, and at least one output device. Program code is applied to input data to perform the functions described herein and generate output information. The output information is applied to one or more output devices, in known fashion.

Each program is preferably implemented in a high level procedural or object oriented programming language to communicate with a computer system. However, the programs can be implemented in assembly or machine language, if desired. In any case, the language may be a compiled or interpreted language.

Each such computer program is preferably stored on a storage media or device (*e.g.*, ROM or magnetic diskette) readable by a general or special purpose programmable computer, for configuring and operating the computer when the storage media or device is read by the computer to perform the procedures described herein. The inventive system may also be considered to be implemented as a

computer-readable storage medium, configured with a computer program, where the storage medium so configured causes a computer to operate in a specific and predefined manner to perform the functions described herein.

A number of embodiments of the present invention have been described. Nevertheless, it will be understood that various modifications may be made without departing from the spirit and scope of the invention. Accordingly, other embodiments are within the scope of the following claims.

CLAIMS

What is claimed is:

1. A method, performed on a computer system, of regulating a user's access to a computer-based service, the method comprising the steps of:
   (a) receiving input about a first user from at least one other user of the computer-based service;
   (b) evaluating the received input; and
   (c) modifying the first user's ability to access the computer-based service based on a result of such evaluating.

2. The method of claim 1, in which the step of receiving input comprises gathering a vote from at least one other user of the computer-based service, the vote being cast in response to action taken by the first user.

3. The method of claim 2, in which the step of evaluating comprises applying predetermined access limiting criteria to the cast vote.

4. The method of claim 2, in which the step of evaluating further includes determining whether each other user is permitted to cast a vote against the first user.

5. The method of claim 4, in which the step of determining further includes evaluating whether a voting user has been accessing at least a portion of the computer-based service for a minimum amount of time.

6. The method of claim 4, in which the step of determining further includes evaluating whether a voting user has been accessing at least a portion of the computer-based service for a longer amount of time than the first user.

7. The method of claim 4, in which the step of determining further includes limiting the number of votes a user may cast.

8. The method of claim 2, in which the step of evaluating further includes giving extra weight to votes cast by another user against the first user if the other user has been accessing at least a portion of the computer-based service for a selected amount of time.

9. The method of claim 2, in which the step of modifying comprises denying the first user full access to the computer-based service based on the result of the evaluating.

10. The method of claim 9, further comprising the step of allowing the first user to regain full access to the computer-based service based on predetermined access resumption criteria.

11. The method of claim 1, further comprising the step of notifying selected other users that the first user's ability to access the computer-based service has been modified.

12. The method of claim 1, further comprising the step of providing an indicator of an amount of modification by all users accessing at least a portion of the computer-based service.

13. The method of claim 1, in which the computer-based service comprises an online public forum.

14. The method of claim 1, further comprising the step of permitting the first user to vote against another user based on a message type sent by such other user after the first user's ability to access the computer-based service has been modified.

15. A method, performed on a computer system, of regulating a user's access to a computer-based service, the method comprising the steps of:

(a) computing an average message rate for messages originated by the user's connection to the computer-based service;

(b) comparing the user's average message rate to a rate limit threshold;

(c) if the average message rate exceeds the rate limit threshold, then modifying the user's ability to access the computer-based service.

16. The method of claim 15, further comprising the steps of:

(a) comparing the user's average message rate to a disconnect threshold;

(b) if the average message rate exceeds the disconnect threshold, then denying the user access to the computer-based service.

17. The method of claim 15, further comprising the step of allowing the user full access to the computer-based service if the average message rate falls below a clear threshold.

18. A method, performed on a computer system, of regulating a user's access to a computer-based service, the method comprising the steps of:

(a) computing an average message rate for messages originated by a user's connection to the computer-based service;

(b) comparing the user's average message rate to a rate limit threshold;

(c) if the average message rate exceeds the rate limit threshold, then modifying the user's ability to access the computer-based service;

(d) comparing the user's average message rate to a disconnect threshold;

(e) if the average message rate exceeds the disconnect threshold, then denying the user access to the computer-based service;

(f) allowing the user full access to the computer-based service if the average message rate falls below a clear threshold.

19. The method of claim 18, further comprising the step of adjusting any of the average message rate, the rate limit threshold, the disconnect threshold, or the clear threshold in response to negative input about the user from at least one other user of the computer-based service.

20. The method of claim 18, wherein the step of computing an average message rate for messages originated by the user's connection includes counting only messages of selected types.

21. The method of claim 18, wherein the step of computing an average message rate for messages includes starting with an initial average message rate.

22. The method of claim 21, further comprising adjusting a value of the initial average message rate to vary the computer-based service's tolerance of initial messaging activity.

23. A computer program, residing on a computer-readable medium, for regulating a user's access to a computer-based service, comprising instructions for causing a computer to:

(a) receive input about a first user from at least one other user of the computer-based service;

(b) evaluate the received input; and

(c) modify the first user's ability to access the computer-based service based on a result of such evaluating.

24. The computer program of claim 21, in which the instructions for causing a computer to receive input further comprises instructions for causing a computer to gather a vote from at least one other user of the computer-based service, the vote being cast in response to action taken by the first user.

25. The computer program of claim 22, in which the instructions for causing a computer to evaluate further comprises instructions for causing a computer to apply predetermined access limiting criteria to the cast vote.

26. The computer program of claim 22, in which the instructions for causing a computer to evaluate further comprises instructions for causing a computer to determine whether each other user is permitted to cast a vote against the first user.

27. The computer program of claim 24, in which the instructions for causing a computer to determine further comprises instructions for causing a computer to evaluate whether a voting user has been accessing at least a portion of the computer-based service for a minimum amount of time.

28. The computer program of claim 24, in which the instructions for causing a computer to determine further comprises instructions for causing a computer to evaluate whether a voting user has been accessing at least a portion of the computer-based service for a longer amount of time than the first user.

29. The computer program of claim 24, in which the instructions for causing a computer to determine further comprises instructions for causing a computer to limit the number of votes a user may cast.

30. The computer program of claim 22, in which the instructions for causing a computer to evaluate further comprises instructions for causing a computer to give extra weight to votes cast by another user against the first user if the other user has been accessing at least a portion of the computer-based service for a selected amount of time.

31. The computer program of claim 22, in which the instructions for causing a computer to modify further comprises instructions for causing a computer to deny the first user full access to the computer-based service based on the result of the evaluation.

32. The computer program of claim 29, further comprising instructions for causing a computer to allow the first user to regain full access to the computer-based service depending on predetermined access resumption criteria.

33. The computer program of claim 21, further comprising instructions for causing a computer to notify selected other users of modification of the first user's ability to access the computer-based service.

34. The computer program of claim 21, further comprising instructions for causing a computer to provide an indicator of an amount of modification by all users accessing at least a portion of the computer-based service.

35. The computer program of claim 21, in which the computer-based service comprises an online public forum.

36. The computer program of claim 21, further comprising instructions for causing a computer to permit the first user to vote against another user based on message type sent by such other user after the first user's ability to access the computer-based service has been modified.

37. A computer program, residing on a computer-readable medium, for regulating a user's access to a computer-based service, comprising instructions for causing a computer to:

(a) compute an average message rate for messages originated by the user connection to the computer-based service;

(b) compare the user's average message rate to a rate limit threshold;

(c) if the average message rate exceeds the rate limit threshold, then modify the user's ability to access the computer-based service.

38. The computer program of claim 35, further comprising instructions for causing a computer to:

(a) compare the user's average message rate to a disconnect threshold;

(b) if the average message rate exceeds the disconnect threshold, then deny the user access to the computer-based service.

39. The computer program of claim 35, further comprising instructions for causing a computer to allow the user full access to the computer-based service if the average message rate falls below a clear threshold.

40. A computer program, residing on a computer-readable medium, for regulating a user's access to a computer-based service, comprising instructions for causing a computer to:

(a) compute an average message rate for messages originated by a user's connection to the computer-based service;

(b) compare the user's average message rate to a rate limit threshold;

(c) if the average message rate exceeds the rate limit threshold, then modify the user's ability to access the computer-based service;

(d) compare the user's average message rate to a disconnect threshold;

(e) if the average message rate exceeds the disconnect threshold, then deny the user access to the computer-based service;

(f) allow the user full access to the computer-based service if the average message rate falls below a clear threshold.

41. The computer program of claim 40, further comprising instructions for causing a computer to adjust any of the average message rate, the rate limit threshold, the disconnect threshold, or the clear threshold in response to negative input about the user from at least one other user of the computer-based service.

42. The computer program of claim 40, in which the instructions for causing a computer to compute an average message rate for messages originated by the user's connection include instructions for causing a computer to count only messages of selected types.

43. The computer program of claim 40, wherein the instructions to compute an average message rate for messages includes starting with an initial average message rate.

44. The computer program of claim 43, further comprising instructions to adjust a value of the initial average message rate to vary the computer-

based service's tolerance of initial messaging activity.

45. A method, performed on a computer system, of monitoring user activity in a computer-based service, the method comprising:
receiving input about a first user's activities from at least one other user of the computer-based service; and
adjusting an index representative of the first user's activities based on the received input.

46. The method of claim 45, further comprising modifying the first user's ability to access the computer-based service based on the first user's index.

47. The method of claim 45, in which the input received from the at least one other user is in response to action taken by the first user.

48. The method of claim 45, further comprising limiting received input only to other users who have been effected by the first user's activities.

49. The method of claim 45, in which the index representative of the first user's activities corresponds to the appropriateness of the first user's behavior in using the computer-based service.

50. The method of claim 45, in which the input received from the at least one other user is sent anonymously.

51. The method of claim 50, in which adjustment of the first user's index comprises according reduced significance to anonymous input received from the at least one other user.

52. The method of claim 45, in which the input received from the at least one other user is sent automatically in response to action taken by the first user.

53. The method of claim 52, in which adjustment of the first user's index comprises according reduced significance to automatic input received from the at least one other user.

54. The method of claim 45, in which adjustment of the first user's index comprises according varied significance to received input based on a class to which the at least one other user belongs.

55. The method of claim 54, in which adjustment of the first user's index comprises according increased significance to input received from a user belonging to a system operator class.

56. The method of claim 54, in which adjustment of the first user's index comprises according reduced significance to input received from a user belonging to a non-paying member class.

57. The method of claim 54, in which adjustment of the first user's index comprises according varied significance to received input based on a seniority level of the at least one other user.

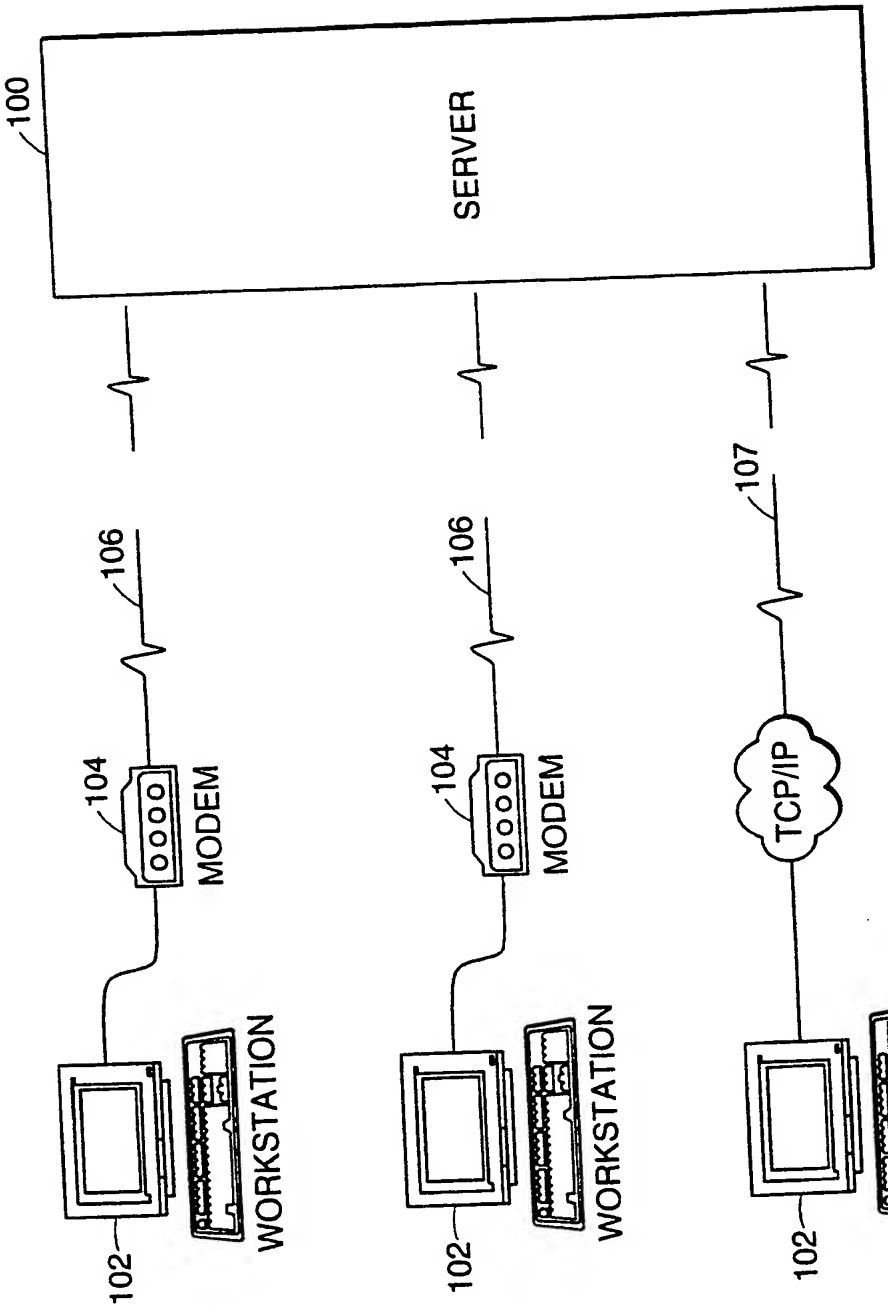


FIG. 1
(PRIOR ART)

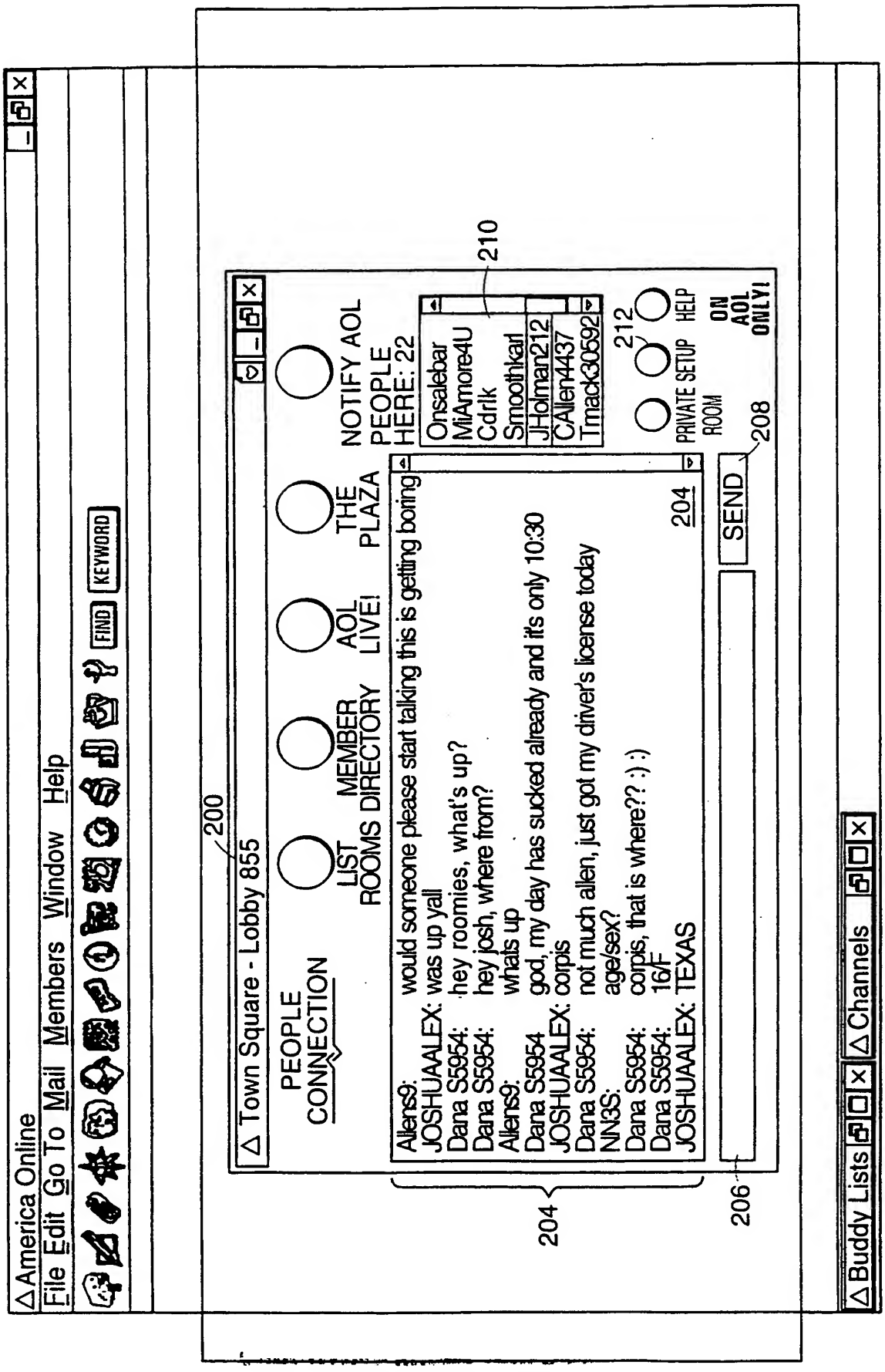


FIG. 2
(PRIOR ART)

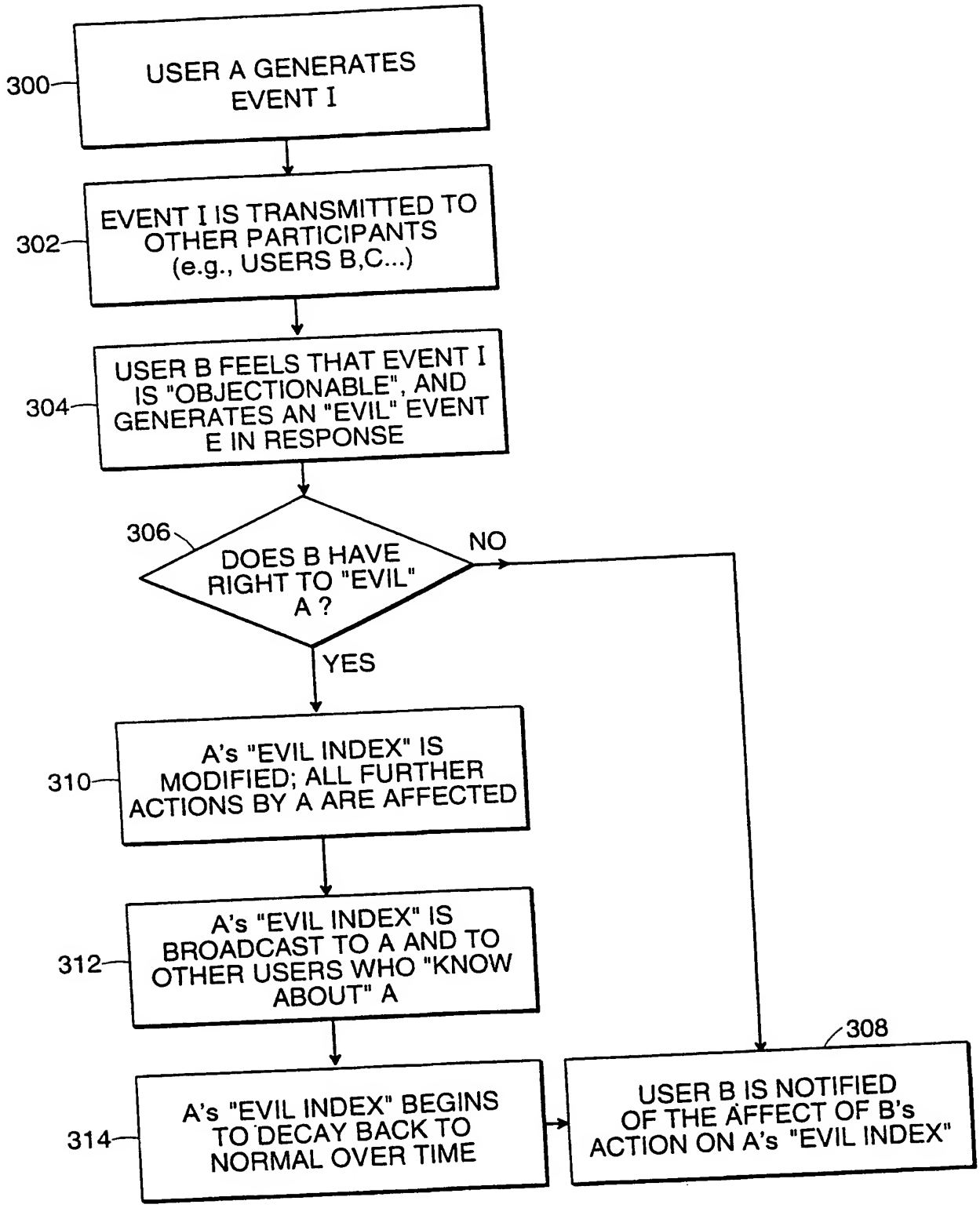

300
USER A GENERATES EVENT I
302
EVENT I IS TRANSMITTED TO OTHER PARTICIPANTS (e.g., USERS B,C...)
304
USER B FEELS THAT EVENT I IS "OBJECTIONABLE", AND GENERATES AN "EVIL" EVENT E IN RESPONSE
306
DOES B HAVE RIGHT TO "EVIL" A ?
NO
YES
310
A's "EVIL INDEX" IS MODIFIED; ALL FURTHER ACTIONS BY A ARE AFFECTED
312
A's "EVIL INDEX" IS BROADCAST TO A AND TO OTHER USERS WHO "KNOW ABOUT" A
314
A's "EVIL INDEX" BEGINS TO DECAY BACK TO NORMAL OVER TIME
308
USER B IS NOTIFIED OF THE AFFECT OF B's ACTION ON A's "EVIL INDEX"

FIG. 3

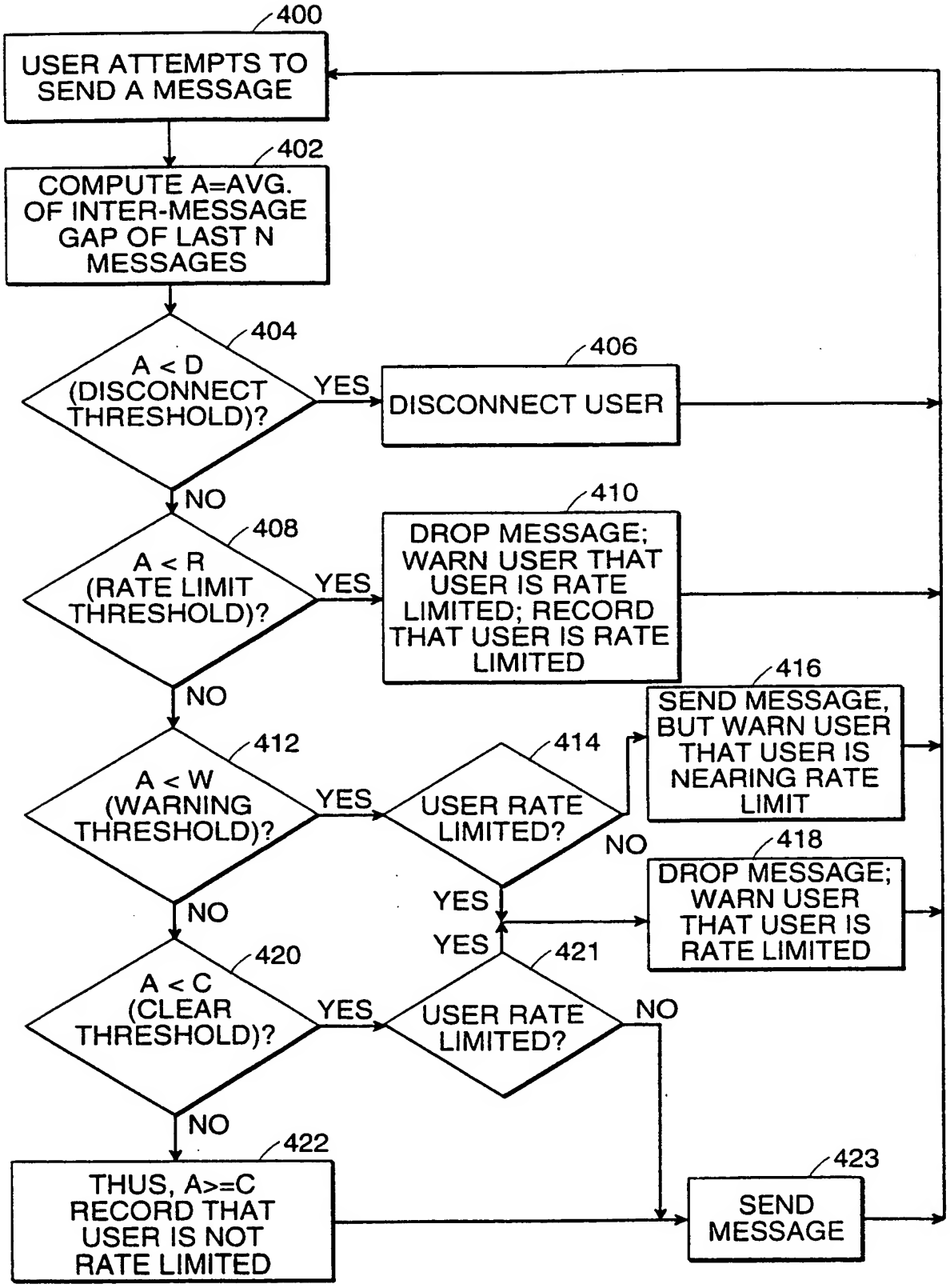

400
USER ATTEMPTS TO SEND A MESSAGE
402
COMPUTE A=AVG. OF INTER-MESSAGE GAP OF LAST N MESSAGES
404
A < D (DISCONNECT THRESHOLD)?
YES
406
DISCONNECT USER
NO
408
A < R (RATE LIMIT THRESHOLD)?
YES
410
DROP MESSAGE; WARN USER THAT USER IS RATE LIMITED; RECORD THAT USER IS RATE LIMITED
NO
412
A < W (WARNING THRESHOLD)?
YES
414
USER RATE LIMITED?
YES
416
SEND MESSAGE, BUT WARN USER THAT USER IS NEARING RATE LIMIT
NO
418
DROP MESSAGE; WARN USER THAT USER IS RATE LIMITED
NO
420
A < C (CLEAR THRESHOLD)?
YES
421
USER RATE LIMITED?
YES
NO
NO
422
THUS, A>=C RECORD THAT USER IS NOT RATE LIMITED
423
SEND MESSAGE

FIG. 4

INTERNATIONAL SEARCH REPORT

Int. ional Application No
PCT/US 98/09969

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G06F1/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	J. BERKMAN: "Nonviolent Crowd Control" PROCEEDINGS OF THE ACM SIGUCCS 1993 USER SERVICES CONFERENCE XXI, 7 - 10 November 1993, SAN DIEGO, CA, US, pages 173-178, XP002074878 see page 174 - page 175 see page 177 ---	1,13,15, 18,20, 23,35, 37,40, 42,45
A	J. JOHNSON-EILOLA ET AL.: "Policing Ourselves: Defining the Boundaries of Appropriate Discussion in Online Forums " COMPUTERS AND COMPOSITION, vol. 13, no. 3, 1996, US, pages 269-291, XP002074883 see page 271 see page 273 see page 275 - page 282 --- -/--	1,13,23, 35

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publicationdate of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of theinternational search	Date of mailing of the international search report
19 August 1998	07/09/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk . Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Taylor, P

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Int. ional Application No
PCT/US 98/09969

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 265 221 A (MILLER) 23 November 1993 see column 1, line 12 - column 3, line 22 see column 5, line 22 - column 6, line 38; figures 2,3 -----	1,13,23, 35,45, 54,55,57

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. ional Application No
PCT/US 98/09969

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5265221 A	23-11-1993	NONE	